

Annual Report
January 3, 2003

Numerical Simulations For Active Tectonic Processes:
Increasing Interoperability And Performance

JPL Task Plan No. 83-6791

Administration Milestone B – 8/30/2002

First Annual Report delivered.

Team

Andrea Donnellan: Principle Investigator, Database design and implementation

Jet Propulsion Laboratory
Mail Stop 183-335
4800 Oak Grove Drive
Pasadena, CA 91109-8099
donnellan@jpl.nasa.gov
818-354-4737

Geoffrey Fox: Information Architect

Community Grid Computing laboratory
Indiana University
501 N. Morton, Suite 224
Bloomington, IN 47404-3730
gcf@indiana.edu
812-856-7977

Jay Parker: Software engineer

Jet Propulsion Laboratory
Mail Stop 238-600
4800 Oak Grove Drive
Pasadena, CA 91109-8099
Jay.W.Parker@jpl.nasa.gov
818-354-6790

Marlon Pierce: Code Interoperability Software Engineer

Community Grid Computing laboratory
Indiana University
501 N. Morton, Suite 224
Bloomington, IN 47404-3730
marpierc@indiana.edu
812-856-1212

Dennis McLeod: Database interoperability

University of Southern California
Mail Code 0781
3651 Trousdale Parkway
Los Angeles, CA 90089-0742
mcloed@pollux.usc.edu
213-740-7285

John Rundle: Pattern Recognizers

University of California, Davis
Department of Physics
One Shields Avenue
Davis, CA 95616-8677
rundle@physics.ucdavis.edu
530-752-6416

Lisa Grant: Fault database architect

University of California, Irvine
Environmental Analysis and Design
Irvine, CA 92697-7070
lgrant@uci.edu
949-824-5491

Terry Tullis: Fast Multipole Methods

Brown University
Department of Geological Sciences
Brown University
Providence, RI 02912-1846
Terry_Tullis@brown.edu
401-863-3829

Robert Granat: Pattern Recognizers

Jet Propulsion Laboratory
Mail Stop 126-347
4800 Oak Grove Drive
Pasadena, CA 91109-8099

Robert.Granat@jpl.nasa.gov
818-393-5353

Gregory Lyzenga: Finite Element Models

Jet Propulsion Laboratory
Mail Stop 126-347
4800 Oak Grove Drive
Pasadena, CA 91109-8099
Gregory.Lyzenga@jpl.nasa.gov
818-354-6920

Objective

The full objective over this three-year program is to produce a system to fully model earthquake-related data. Components of this system include:

- A database system for handling both real and simulated data
- Fully three-dimensional finite element code (FEM) with adaptive mesh generator capable of running on workstations and supercomputers for carrying out earthquake simulations
- Inversion algorithms and assimilation codes for constraining the models and simulations with data
- A collaborative portal (object broker) for allowing for seamless communication between codes, reference models, and data
- Visualization codes for interpretation of data and models
- Pattern recognizers capable of running on workstations and supercomputers for analyzing data and simulations

In order to develop a solid earth science framework for understanding and studying of active tectonic and earthquake processes, this task develops simulation and analysis tools to study the physics of earthquakes using state-of-the-art modeling, data manipulation, and pattern recognition technologies. We develop clearly defined accessible data formats and code protocols as inputs to the simulations. These are adapted to high-performance computers because the solid earth system is extremely complex and nonlinear resulting in computationally intensive problems with millions of unknowns. With these tools it will be possible to construct the more complex models and simulations necessary to develop hazard assessment systems critical for reducing future losses from major earthquakes.

Approach

This task combines several areas of computational science and geophysics and involves experts at 6 institutions. The components of the framework, the team members developing them, and the tasks, backgrounds and skills are as follows.

Component	Team Members	Task, Background, and Skills
Database system	Dennis McLeod Lisa Grant Andrea Donnellan	Database systems Geology/paleoseismology Tectonics, earthquakes
Finite element code	Greg Lyzenga Jay Parker John Lou	Parallel computation/FEM Parallel computation/FEM Parallel computation/FEM/AMR
Inversion and assimilation codes	Greg Lyzenga John Rundle Terry Tullis John Lou	Inversion Viscoelastic models Fault nucleation, fast multipoles Fast multipoles, data assimilation
Web Services	Geoffrey Fox Marlon Pierce	Science interoperability Science interoperability
Visualization codes	Geoffrey Fox Andrea Donnellan	Visualization (using students shared with FSU visualization group) and integration in portals Geophysical applications
Pattern recognizers	Robert Granat John Rundle	Hidden Markov Modeling Pattern dynamics

This table with additional information is posted at <http://www-aig.jpl.nasa.gov/public/dus/quakesim/people.html>.

Milestones pertaining to the full interoperable system are subjected to the approval of an independent review board, with members:

- Brad Hager, MIT (Chair)
- Gordon Erlebacher, Florida State University
- Louise Kellogg, UC Davis
- Julia K. Morgan, Rice University
- Gilles Peltzer, UCLA/JPL
- Mark Simons, Caltech
- Don Turcotte, University of California, Davis

We are building a new Problem Solving Environment (QuakeSim) for use by the seismological, crustal deformation, and tectonics communities for developing an understanding of active tectonic and earthquake processes. The top-level operational architecture of our proposed solid earth research virtual observatory (SERVO) shows science users interacting with interface programs as well as modeling, simulation, and analysis tools. The general architecture follows the “Web Services” model being developed by business interests, but is applied to scientific applications and supporting software resources (such as databases). The system is divided into three tiers: a user interface layer (implemented as a browser interface), a system resource layer, and a middle control layer that maintains proxies (or brokers) to the system resources (Figure 1). The middle tier provides a uniform interface to the resource layer. Following the Web Services approach, we define XML interface abstractions (in WSDL) for basic

services (such as File Management) and implement the interface with appropriate technologies (such as with a relational database). Communication between the services is done with an XML messaging architecture (SOAP).

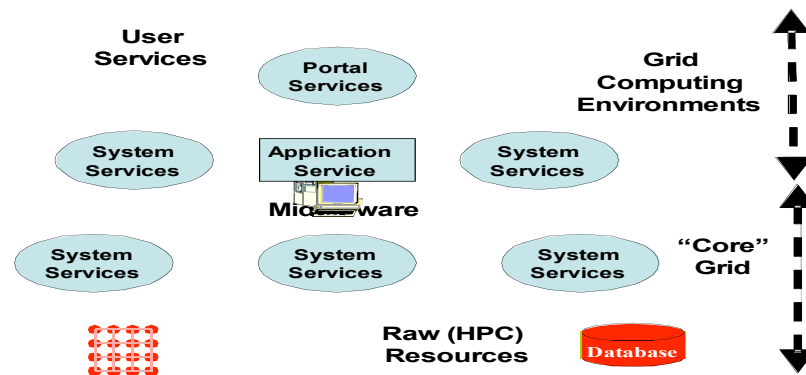


Figure 1. High-level architecture of planned system showing grids, portals, and grid computing environments.

One of the most critical aspects of our proposed system is supporting interoperability given the heterogeneous nature of data sources as well as the variety of application programs, tools, and simulation packages that must operate with data from our system. Interoperability will be implemented by using distributed object technology combined with development of object API's that conform to emerging standards. We will define our object API's in XML and dynamically map this specification into the chosen object model. This strategy was successfully used in the Gateway portal, which currently uses a CORBA middle tier but has used a pure Java solution with the same objects.

The preliminary design for this interoperability framework has been approved and is posted at <http://www-aig.jpl.nasa.gov/public/dus/quakesim/DesignDocument3.0.doc>. Several codes have been identified to be part of the interoperable system and carry out functions to calculate surface displacements and slip history due to fault slip or carry out pattern recognition on real or simulated data.

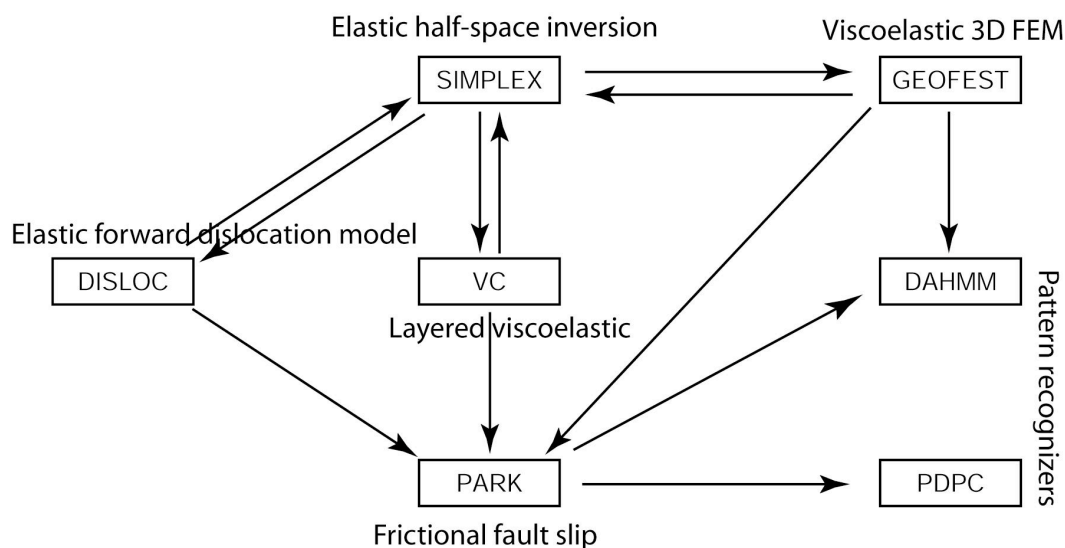


Figure 2. Linkages between programs that maybe run separately or coupled to other programs. Low-level movement of actual data occurs between programs. There is also communication of metadata and information about the files being transferred.

Three of these codes are targeted for improved performance, chiefly through design changes that make them efficient high-performance parallel codes. These codes are PARK, a boundary-element based code for studying unstable slip at the Parkfield segment of the San Andreas fault, Virtual California, which simulates the dynamic interaction of hundreds of fault segments comprising the active tectonics of California, and GeoFEST, a fully three-dimensional finite element code to model active tectonics and earthquake processes. Together with an adaptive mesh generator that constructs a mesh based on geometric and mechanical properties of the crustal structure, the GeoFEST system makes it possible to efficiently model time-dependent deformation of interacting fault systems embedded in a heterogeneous earth structure.

Scientific Accomplishments

As this report reflects the first few months of this task, the scientific accomplishments are best represented by a discussion of the expected utility of the system of applications we are developing. However, several investigators are simultaneously carrying out scientific research and reflected in the talks and papers listed below.

The full interoperable system will allow users from many environments to discover and exploit a very wide range of applications, models, and physical measurements in distributed databases. This is seen as of crucial importance for gaining full advantage of massive new national programs for gathering new regional and global data sets, such as the Plate Boundary Observatory, Earthscope, and NASA orbiting inSAR missions.

PARK is a model for unstable slip on a single earthquake fault. Because it aims to capture the instability it is designed to represent the slip on a fault at many scales, and to capture the developing seismic slip details over an extraordinary range of time scales

(subseconds to decades). Its simulation of the evolution of fault rupture is the most realistic of the tools considered here. When transformed into an efficient parallel simulation, it will be the tool of choice for researchers seeking to determine the nature and detectability of earthquake warning signals such as surface strains and patterns of microseismicity. This is the first earthquake simulation code to seek enhanced scalability and speed by employing a multipole technique. The multipole experience gained here will also be transferable to the Virtual California code and other boundary element simulations. The power of massive parallel computing is required for this problem in order to support many small slip patch elements in order to cover the nucleation scale that initiates the instability.

GeoFEST simulates stress evolution, fault slip and plastic/elastic processes in realistic materials. The products of such simulations are synthetic observable time-dependent surface deformation on scales from days to decades. Diverse types of synthetic observations will enable a wide range of data assimilation and inversion techniques for ferreting out subsurface structure and stress history. In the short term, such a tool allows rigorous comparisons of competing models for interseismic stress evolution, and the sequential GeoFEST system is being used for this at JPL and UC Davis. Parallel implementation is required to go from local, single-event models to regional models that cover many earthquake events and cycles.

Virtual California simulates fault interaction to determine correlated patterns in the nonlinear complex system of an entire plate boundary region. The evolution of these patterns enables forecasts of future large events. The model produces synthetic seismicity and surface deformation, enabling an eventual data assimilation system for exploiting regional data collection. Capturing the nonlinear pattern dynamics of the fault system along a plate boundary implies the realization of a digital laboratory, which allows understanding of the mechanisms behind the observations and patterns. Our technology development aims to produce and demonstrate a scalable cluster code. When that is deployed researchers will be able to create and verify patterns down to smaller spatial scales, which will enable cross-scale parameterization and validations, which will in turn enable plate-boundary system analysis and greatly enhanced forecasts of large earthquakes.

Pattern analysis methods are another tool type we are developing. One method (Rundle et al., 2002, see cover graphic) bins many decades of seismic activity on a gridded representation of California. Eigensystem analysis reveals clusters of correlated space and time activity, which have been subjected to a phase dynamics forecast. When this method attains parallel speedup we will produce better forecasts and enable speedy tests of earthquake correlations and predictions. This will be due to the ability to use much smaller geographic cell sizes and so forecast the frequent magnitude 4 earthquakes, not just the rare magnitude 6 events.

Technology Accomplishments

The Milestone E code baseline benchmark marks the beginning of moving the three codes, PARK, geoFEST, and Virtual California to high-performance parallel computers, enabling the simulation of interacting fault systems leading to earthquakes. The scientific

significance and performance results have been reported in http://www-aig.jpl.nasa.gov/public/dus/quakesim/MilestoneE_Baselines.pdf and are summarized in the following table:

Code	Resolution	Time Steps	Platform (processors)	Wallclock time
PARK	15,000 patches	500 steps	SGI Origin 3000 (1)	7h53m
geoFEST	50,000 elements	1000 steps	Sun workstation (1)	13h43m
Virtual California	215 segments	10,000 steps	Pentium III (1)	5m38s (stress GF) 15m 48s (steps)

Performance for PARK on a single processor can handle a fault region with 15,000 fault elements over 500 time steps in just under 8 hours. But the problem of determining possible earthquake precursors requires a finer sampling. This project is committed to demonstrating the computation of 400,000 elements on 1024 processors, enabling 50,000 time steps in 40 hours by June 2004. Such performance will be attained by combining efficient parallel problem decomposition and by exploiting an NlogN multipole technique to reduce the currently N^2 element interaction computations. This multipole method has been demonstrated in parallel astrophysics simulations with considerable success.

Performance for Virtual California has been demonstrated at 215 interacting fault segments for 10,000 time steps in 20 minutes on a 1 GHz Linux workstation. This type of simulation allows determination of correlated patterns of activity that can be used to forecast seismic hazard, with relevance for earthquakes of Magnitude 6. Finer sampling may allow verification of such forecasts by including much smaller, more frequent earthquakes. The aims of this task are to validate parallel scaling of the Virtual California code, and evaluate the potential for the multipole technique pioneered in the PARK demonstration.

GeoFEST is a direct implementation of stress-displacement volumetric finite elements. The mechanics include elasticity and visco-elastic deformation. Current workstation capacity can solve for the deformation and stress evolution due to a single-fault rupture, such as the Northridge event, using $\sim 100,000$ finite element equations. We seek to analyze the modes of interaction of the entire Southern California system of interacting faults, covering a portion of the crust ~ 1000 km on a side. Such a simulation would require $\sim 5M$ equations to determine first-order effects, and certainly higher density for faithful representation of the time-dependent stress field. Current techniques require running such a model through thousands of time steps to attain a stable background stress field and assess the patterns of fault interactions. These considerations motivate tailoring the code toward hundreds of processors to attain solutions in reasonable turnaround time.

Our team has created a meshing tool with mesh generation and adaptation capabilities, while the ESS project has developed a parallel adaptive meshing library (the Pyramid library) for supporting parallel dynamic adaptive meshing of unstructured meshes. Techniques of parallel adaptive meshing, combined with local error estimates, can be

effectively used to compute an initial state of the displacement and stress fields, and to significantly reduce the computational load of computing the evolving stress field in the finite element modeling.

We are therefore in an excellent position to evaluate the ESS Adaptive Mesh Refinement (AMR) library with our finite element code, and to combine the strengths of both meshing tools into an integrated adaptive meshing library for a scalable parallel computing environment. The integrated adaptive meshing library will support the entire process of an unstructured parallel application, including initial mesh generation, dynamic parallel mesh adaptation with efficient mesh quality control, and dynamic load balancing. This will be a valuable tool for improving the computational efficiency of our finite element modeling.

Current GeoFEST performance computes 1000 time steps, and 50,000 volumetric finite elements with an unstructured mesh with large range of element sizes, in just under 13 hours on a 400 Mhz Sun Solaris computer. We plan to demonstrate problems with a complex embedded system of faults with 16,000,000 finite elements with the same number of time steps and solution time on 880 processors. We will use PYRAMID for parallel decomposition and adaptive mesh refinement, and a parallel iterative Conjugate Gradient method for the sparse matrix solution.

The Milestone H interoperability milestone established a preliminary design for the framework, which will enable seamless access to data and models across the Internet.

Status/Plans

In this fractional first year we have

1. Agreed upon a software engineering plan,
 - Selected a review board of regarded earthquake and computational science experts,
 - Established a public web page area (<http://www-aig.jpl.nasa.gov/public/dus/quakesim/>) to allow documentation and access to our products,
 - Demonstrated and documented the serial performance of three important codes believed to be scalable to high performance computing.
 - Established requirements and design policy for a framework that will make community earthquake codes interoperable and deliverable to the scientific community.

In the coming year we will develop the three main codes for efficient parallel performance. Virtual California and geoFEST will be adapted for Riva and Parvox parallel rendering. Also, geoFEST will be linked with the PYRAMID mesh refinement library. We will develop the interoperability framework and tie in a fault database, mesh generator and the geoFEST code and demonstrate the prototype framework with a combined simulation.

Point of contact

Andrea Donnellan
Jet Propulsion Laboratory
Mail Stop 183-335
4800 Oak Grove Drive
Pasadena, CA 91109-8099
donnellan@jpl.nasa.gov
818-354-4737

Caption for Cover Graphic

Increasing log probability (forecast) of large seismic events for the decade 2000-2010 based on changes in seismic patterns during 1989-1999 according to phase dynamics method (Rundle et al., 2002); inverted triangles mark large events in 1989-1999, circles mark large events since 2000. Zoom view of same forecast in vicinity of Northridge 1994 quake with black dots showing contour of post-seismic motion. InSAR fringes and GPS motion (arrows) for two-year postseismic period with yellow dots showing contour of the region of postseismic deformation. Synthetic view of Northridge region from northwest combining Landsat shading, 5X-exaggerated digital elevation model relief, and seismic plus 27-year coseismic relaxation uplift from 3-layer finite element model shown as brighter colors.

Publications

Donnellan, A., J. Parker, and G. Peltzer, Combined GPS and InSAR models of postseismic deformation from the Northridge earthquake, *Pure and Appl. Geophys.*, 2261–2270, 2002.

Donnellan, Andrea, Geoffrey Fox, John Rundle, Dennis McLeod, Terry Tullis, Lisa Grant, Jay Parker, Marlon Pierce, Gregory Lyzenga, Anne Chen, John Lou, The Solid Earth Research Virtual Observatory: A web-based system for modeling multi-scale earthquake processes, *EOS Trans, AGU*, 83(47), Fall Meeting Suppl., Abstract NG528-08, Invited, 2002.

Fox, Geoffrey, Hasan Bulut, Kangseok Kim, Sung-Hoon Ko, Sangmi Lee, Sangyoon Oh, Shrideep Pallickara, Xiaohong Qiu, Ahmet Uyar, Minjun Wang, Wenjun Wu, Collaborative Web Services and Peer-to-Peer Grids to be presented at 2003 Collaborative Technologies Symposium (CTS'03).

<http://grids.ucs.indiana.edu/ptliupages/publications/foxwmc03keynote.pdf>

- Fox, Geoffrey, Hasan Bulut, Kangseok Kim, Sung-Hoon Ko, Sangmi Lee, Sangyoon Oh, Xi Rao, Shrideep Pallickara, Quinlin Pei, Marlon Pierce, Ahmet Uyar, Wenjun Wu, Choonhan Youn, Dennis Gannon, and Aleksander Slominski, "An Architecture for e-Science and its Implications" in *Proceedings of the 2002 International Symposium on Performance Evaluation of Computer and Telecommunications Systems*, edited by Mohammed S.Obaidat, Franco Davoli, Ibrahim Onyuksel and Raffaele Bolla, Society for Modeling and Simulation International, pp 14-24 (2002).
<http://grids.ucs.indiana.edu/ptliupages/publications/spectscience.pdf>
- Fox, Geoffrey, Sung-Hoon Ko, Marlon Pierce, Ozgur Balsoy, Jake Kim, Sangmi Lee, Kangseok Kim, Sangyoon Oh, Xi Rao, Mustafa Varank, Hasan Bulut, Gurhan Gunduz, Xiaohong Qiu, Shrideep Pallickara, Ahmet Uyar, Choonhan Youn, *Grid Services for Earthquake Science, Concurrency and Computation: Practice and Experience in ACES Special Issue*, 14, 371-393, 2002.
<http://aspen.ucs.indiana.edu/gemmauisummer2001/resources/gemandit7.doc>
- Gould, M., Grant, L., Donnellan, A., and D. McLeod. The GEM Fault Database: A Preliminary report on design and approach. *Proceedings and Abstracts, SCEC Annual Meeting*, p.75-76, 2002.
- Granat, R., and A. Donnellan, A hidden Markov model tool for geophysical data exploration, *Pure and Appl. Geophys.*, 2271–2284, 2002.
- Grant, L. B. and Gould, M. M. Paleoseismic and geologic data for earthquake simulation. *Computational Science, Data Assimilation, and Information Technology for Understanding Earthquake Physics and Dynamics*, 3rd ACES International Workshop, Maui, Hawaii, USA, p.30, 2002.
- Grant, L. B. and M. M. Gould. Assimilation of paleoseismic data for earthquake simulation. Submitted to *Pure and Applied Geophysics*, 2002.
- Parker, Jay, Gregory Lyzenga, Jin-Fa Lee, John Lou, and Andrea Donnellan, Technologies for larger crustal studies by finite element simulation, Southern California Earthquake Center Annual Meeting, Oxnard, CA, September 8-11 2002.
- Rundle, JB, PB Rundle, W Klein, J Martins, KF Tiampo, A Donnellan and LH Kellogg, GEM plate boundary simulations for the Plate Boundary Observatory: Understanding the physics of earthquakes on complex fault systems, *Pure and Appl. Geophys.*, 159, 2357-2381 2002.
- Rundle, JB, W Klein, KF Tiampo, Andrea Donnellan, and GC Fox, Detection and Analysis of Space-Time Patterns in Complex Driven Threshold Systems, paper submitted to the 2003 International Conference on Computational Science, Melbourne, AU June 2-4, 2003.
- Tiampo, KF, JB Rundle, S. Gross and S. McGinnis, Parallelization of a large-scale computational earthquake simulation program, *Concurrency & Computation, Practice and Experience*, 14, 531-550, ACES Special Issue, 2002.
- Tiampo, KF, JB Rundle, S. McGinnis and W. Klein, Pattern dynamics and forecast methods in seismically active regions, *Pure and Appl. Geophys.*, 159, 2429-2467, 2002.

Media references

Northridge Quake Activity Has Apparently Subsided, Says NASA, JPL press release 2002-183, September 30, 2002. Included B-roll and several live interviews with Andrea Donnellan and Gregory Lyzenga including discussion of this CT project.

Academic Training

Maggi Glasscoe – University of California, Davis, PhD graduate student is using geoFEST to model observed GPS deformation and coupling between the crust and the mantle in the Los Angeles basin.

Teresa Baker – Massachusetts Institute of Technology, BS undergraduate student, is using simplex, disloc, and geoFEST to model deformation associated with the Northridge earthquake and adjacent Ventura basin.

Anne Chen – University of Southern California, PhD graduate student is undertaking development of the fault and federated databases.

Miryha Gould – University of California, Irvine, PhD graduate student is developing the geological aspects of the fault database.